Computer Networks 53 (2009) 279–298

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Remote detection of bottleneck links using spectral and statistical methods

Xinming He [a,*], Christos Papadopoulos [b], John Heidemann [c], Urbashi Mitra [d], Usman Riaz [d]

[a] Cisco Systems Inc., MRBU, 725 Alder Drive, Milpitas, CA 95035, United States
[b] Colorado State University, Computer Science Department, Fort Collins, CO 80523, United States
[c] University of Southern California, Computer Science Department, Los Angeles, CA 90089, United States
[d] University of Southern California, Electrical Engineering Department, Los Angeles, CA 90089, United States

ARTICLE INFO

Article history:
Received 7 May 2007
Received in revised form 26 September 2008
Accepted 3 October 2008
Available online 17 October 2008

Responsible Editor: A. Pitsillides

Keywords:
Spectral analysis
Bottleneck detection
Traffic analysis

ABSTRACT

Persistently saturated links are abnormal conditions that indicate bottlenecks in Internet traffic. Network operators are interested in detecting such links for troubleshooting, to improve capacity planning and traffic estimation, and to detect denial-of-service attacks. Currently bottleneck links can be detected either locally, through SNMP information, or remotely, through active probing or passive flow-based analysis. However, local SNMP information may not be available due to administrative restrictions, and existing remote approaches are not used systematically because of their network or computation overhead. This paper proposes a new approach to remotely detect the presence of bottleneck links using spectral and statistical analysis of traffic. Our approach is *passive*, operates on *aggregate traffic* without flow separation, and supports *remote detection* of bottlenecks, addressing some of the major limitations of existing approaches. Our technique assumes that traffic through the bottleneck is dominated by packets with a common size (typically the maximum transfer unit, for reasons discussed in Section 5.1). With this assumption, we observe that bottlenecks imprint periodicities on packet transmissions based on the packet size and link bandwidth. Such periodicities manifest themselves as strong frequencies in the spectral representation of the aggregate traffic observed at a downstream monitoring point. We propose a detection algorithm based on rigorous statistical methods to detect the presence of bottleneck links by examining strong frequencies in aggregate traffic. We use data from live Internet traces to evaluate the performance of our algorithm under various network conditions. Results show that with proper parameters our algorithm can provide excellent accuracy (up to 95%) even if the traffic through the bottleneck link accounts for less than 10% of the aggregate traffic.

1. Introduction

A link is *saturated* when the offered load at the link exceeds its capacity. A saturated link is very likely to be the *bottleneck* link for traffic passing through it. Links may become saturated for brief moments during normal operation in the Internet. However, with the exception of expensive, highly utilized links (e.g., satellite or deep space links), a *sustained*, saturated link typically implies an abnormal condition in the network. Examples of such links that are of interest to network operators include: (a) an under-provisioned link that may spend most of its time saturated as many users share its capacity; (b) an under-provisioned access link that may indicate a customer's need to purchase more bandwidth; (c) a denial-of-service (DoS) attack that saturates a link near the victim; (d) individual attackers in a distributed DoS attack that saturate their access links as they try to send packets as fast as possible; and (e) links that may become accidentally saturated due to faulty software or hardware.

* Corresponding author. Tel.: +1 213 327 4298.
E-mail addresses: xhe@cisco.com (X. He), christos@cs.colostate.edu (C. Papadopoulos), johnh@isi.edu (J. Heidemann), ubli@usc.edu (U. Mitra), uriaz@usc.edu (U. Riaz).

1389-1286/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.comnet.2008.10.001

These examples are of interest for several reasons. First, information about saturated links is necessary to influence long-term decisions such as capacity planning and traffic matrix estimation, and in short-term response to external attacks or internal bugs. More importantly, bottlenecks represent a performance problem for users of the network. Network operators would like to systematically monitor bottlenecks and report on them to their users. If traffic is limited by an access link, it shows that a user needs to purchase greater capacity. If within the ISP, a bottleneck link represents a problem that may affect multiple users and must be diagnosed. If outside the ISP, the bottleneck link demonstrates that the problem is external.

Currently, saturated links can be detected through direct observation, e.g., network monitoring with SNMP. While network monitoring tools are widely used, they are not a panacea, particularly because access to SNMP data is often administratively limited and does not provide visibility into a customer's behavior or an external network's performance. In addition, SNMP reports are typically averaged over long intervals (5 minutes or more) and so they miss short but recurring saturation events. Finally, in some cases SNMP data may not be collected or processed because of economic or bandwidth costs.

Detecting bottlenecks can also be done remotely by active probing [22,12] or per-flow analysis and traffic correlation [21]. However, such techniques either require additional traffic to be inserted into the network exacerbating the congestion on the bottleneck link, or incur high computational cost as packets have to be separated according to flows and then correlated to detect sharing of bottlenecks. Ideally a network operator would like to install a network monitoring system that can detect saturated links quickly by looking at the aggregate traffic level, and then resort to detailed per-flow analysis only if a problem is detected.

We propose an approach that can remotely detect the presence of bottleneck links in *aggregate* traffic without flow separation. Note that we are interested in *transient* bottlenecks, which are hard to detect with standard methods such as SNMP monitoring Our key observation is that when links are saturated, they send packets out as fast as possible, resulting in regular back-to-back packet transmissions. We call this regular, back-to-back packet stream *bottleneck traffic* as it is rate-limited by the capacity of the saturated (bottleneck) link. If we observe the bottleneck traffic in the frequency domain, the back-to-back packet transmissions exhibit strong periodicities regulated by the bottleneck link capacity and the packet size. While there are potentially an infinite number of combinations of link speed and packet size, resulting in many potential bottleneck frequencies, in practice, both are constrained to a relatively few, common values. Links typically come in discrete capacities (e.g., 1.5 Mbps, 10 Mbps, 45 Mbps, 100 Mbps, 1 Gbps, etc) corresponding to WAN technologies. Packet sizes exhibit strong modes governed by protocol design, including 60 B for TCP acknowledgments, 572 B for the minimum supported Internet datagram, 1500 B for maximum size Ethernet segment. Moreover, transient bottlenecks often result from large file transfers; such transfers use the maximum available packet size. Throughout this paper we assume that the bottlenecks are caused by large flows, dominated by packets with a common size (typically near the maximum transfer unit, for reasons discussed in Section 5.1). We observe that the strong periodicities in the packet transmissions along the bottleneck link can manifest themselves as strong frequencies in the spectral representation of the aggregate traffic observed at a downstream monitoring point, and we show that bottlenecks can be detected despite interference and noise created by irregular packet transmissions of other flows. Thus, we can detect the presence of bottleneck links by detecting the existence of bottleneck traffic in the aggregate traffic in the spectral domain. Our approach builds on top of prior work of spectral analysis of network traffic [2,28,7,14].

The main contribution of this paper is the development of a novel approach to detect bottleneck links through the periodic packet transmissions in network traffic. Our approach is completely passive and incurs no additional network overhead. It can detect the presence of bottleneck links without flow separation even if the traffic through the bottleneck link accounts for less than 10% of the aggregate traffic at the monitoring point. We also investigate the sensitivity of our approach under different network conditions such as different bottleneck bandwidths. We have not investigated deeply the performance of our algorithms when only partial bottleneck traffic is observed. Such scenarios reduce the bottleneck signal and make detection harder. Further investigation is part of our future work.

The rest of the paper is organized as follows. First, we describe potential applications of bottleneck link detection in Section 2 and give an overview of our detection system in Section 3. Then, in Section 4 we present our technique for calculating spectral representation of network traffic. After applying this technique to visually demonstrate spectral characteristics of bottleneck links in Section 5, we propose an automatic detection algorithm based on Maximum Likelihood Detection in Section 6 and evaluate its performance using real Internet traffic in Section 7. Finally we conclude the paper in Section 9.

2. Applications of bottleneck detection

We are not the first to explore the problem of detecting bottlenecks by passively monitoring traffic. Others have used the regularities in the packet transmission along the bottleneck link to detect bottleneck sharing among multiple flows [21]. However, network operators today rarely explore traffic at this level for at least two reasons. First, current approaches to bottleneck traffic detection are computationally expensive, requiring splitting traffic into flows and then combining different flows into groups sharing the same bottleneck, an expense we avoid in this paper. Second, perhaps because of this expense, there has been limited exploration of how useful early detection of bottleneck traffic might be in network operations. We explore several possibilities below.

2.1. Capacity planning and traffic engineering

In capacity planning and traffic engineering it is important to understand which parts of the network are bottle-